

Original

## Validación de una escala reducida de capacidad funcional del cuestionario WOMAC

Antonio Escobar<sup>a,b,\*</sup>, Kalliopi Vrotsou<sup>a</sup>, Amaia Bilbao<sup>b,c</sup>, Jose M.<sup>a</sup> Quintana<sup>b,d</sup>, Lidia García Pérez<sup>b,e</sup> y Carmen Herrera-Espiñeira<sup>f</sup>

<sup>a</sup> Unidad de Investigación, Hospital de Basurto, Bilbao, España

<sup>b</sup> CIBER de Epidemiología y Salud Pública (CIBERESP), España

<sup>c</sup> Fundación Vasca de Innovación e Investigación Sanitarias (BIOEF), Sondika, Bizkaia, España

<sup>d</sup> Unidad de Investigación, Hospital de Galdakao, Galdakao, Bizkaia, España

<sup>e</sup> Servicio de Evaluación y Planificación, Servicio Canario de Salud, España

<sup>f</sup> Hospital Virgen de las Nieves, Granada, España

### INFORMACIÓN DEL ARTÍCULO

#### Historia del artículo:

Recibido el 9 de marzo de 2011

Aceptado el 8 de junio de 2011

On-line el 25 de septiembre de 2011

#### Palabras clave:

WOMAC

Validación

Forma reducida

Remplazo articular

Aplicabilidad

### R E S U M E N

**Objetivos:** El objetivo del estudio fue evaluar la validez, la fiabilidad y la sensibilidad al cambio de la versión española de la dimensión reducida de 7 ítems de capacidad funcional del Western Ontario McMaster Universities Osteoarthritis Index (WOMAC).

**Métodos:** Estudio prospectivo con seguimiento de 1 año realizado con 459 pacientes en lista de espera para prótesis de cadera o rodilla. Se enviaron los cuestionarios WOMAC y EQ-5D en cuatro tiempos: basal, 3, 6 y 12 meses postintervención. Se realizó el proceso de validación, análisis Rasch y acuerdo medido por el gráfico de Bland-Altman.

**Resultados:** Los valores de los coeficientes rho de Spearman y de concordancia de Lin entre la versión completa y la reducida estuvieron entre 0,96 y 0,98. Se encontraron diferencias estadísticamente significativas ( $p < 0,001$ ) entre los tres grupos de gravedad medidos por una pregunta del EQ-5D. Ambos gráficos de Bland-Altman mostraron que no había diferencias sistemáticas y que las diferencias entre ambas escalas no dependen de la gravedad. La escala reducida se ajusta al modelo Rasch. Los valores  $\alpha$  de Cronbach fueron superiores a 0,85. El estudio de la sensibilidad al cambio dio valores superiores a 1,27 en la prótesis total de rodilla y a 1,72 en la de cadera.

**Conclusiones:** La versión española de la escala reducida de capacidad funcional del WOMAC es válida, fiable y sensible al cambio en los pacientes con remplazo articular. Su uso disminuirá la carga de los pacientes, tanto en estudios como en la práctica clínica.

© 2011 SESPAS. Publicado por Elsevier España, S.L. Todos los derechos reservados.

### Validation of a short form of the function dimension of the WOMAC questionnaire

#### A B S T R A C T

**Objective:** The aim of this prospective study was to evaluate the validity, reliability and responsiveness of the Spanish version of the 7-item short-form of the function dimension of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).

**Methods:** We conducted a prospective 1-year cohort study that included 459 patients on the waiting list for total knee or hip replacement. The WOMAC and EQ-5D questionnaires were sent at four time points: baseline and at 3, 6 and 12 months post-surgery. Based on a previously shortened scale, we performed a validation process with Rasch analysis and agreement measured by Bland-Altman plots.

**Results:** The values of Spearman's rho and Lin concordance coefficients between the full and short-form ranged from 0.96 to 0.98. Differences were statistically significant ( $p < 0.001$ ) among the three groups of severity measured by a item in the EQ-5D. The two Bland-Altman plots showed no systematic differences and agreement was not dependent on the severity score. The short form fitted the Rasch model. Cronbach's  $\alpha$  coefficients were higher than 0.85. Responsiveness was higher than 1.27 and 1.72 in knee and hip replacement, respectively.

**Conclusions:** The Spanish 7-item short form of the WOMAC function dimension is valid, reliable and responsive for patients with total joint replacement. The use of this version will reduce the burden on patients in both clinical studies and clinical practice.

© 2011 SESPAS. Published by Elsevier España, S.L. All rights reserved.

#### Keywords:

WOMAC

Validation

Short form

Total joint replacement

Applicability

\* Autor para correspondencia.

Correo electrónico: antonio.escobarmartinez@osakidetza.net (A. Escobar).

## Introducción

La artrosis de cadera o de rodilla es una de las principales causas de dolor y limitación de la capacidad funcional. Las artroplastias son intervenciones dirigidas a aliviar el dolor y mejorar la capacidad funcional y la calidad de vida relacionada con la salud (CVRS) de estos pacientes<sup>1-3</sup>.

La valoración de la CVRS es una prioridad en las enfermedades crónicas, en las cuales el objetivo del tratamiento es mejorar el estado funcional y la CVRS. En los pacientes con artrosis de miembros inferiores se han utilizado diferentes instrumentos específicos para medir el dolor, la capacidad funcional y la CVRS. Uno de los más utilizados es el Western Ontario McMaster Universities Osteoarthritis Index (WOMAC)<sup>1</sup>, que es uno de los mejores cuestionarios desde el punto de vista de sus propiedades psicométricas<sup>4</sup>. Este cuestionario tiene tres dimensiones que miden el dolor, la rigidez y la capacidad funcional<sup>5</sup>. El cuestionario original demostró tener unas adecuadas propiedades de validez, fiabilidad y sensibilidad al cambio<sup>6</sup>. Por otro lado, se recomienda la inclusión de la subescala de capacidad funcional en ensayos clínicos en pacientes con artrosis de cadera o rodilla<sup>7</sup>. También se recomienda su uso junto a un cuestionario genérico, el Short-Form-36, en la evaluación de los pacientes con artroplastia de miembros inferiores<sup>8</sup>.

La dimensión de capacidad funcional es la más larga de las tres que incluye el WOMAC (consta de 17 ítems). Una versión reducida de esta dimensión que conservase las propiedades psicométricas sería útil, tanto en la práctica clínica como en estudios epidemiológicos, al disminuir la carga de los pacientes, sobre todo en estudios prospectivos en que el paciente debe cumplimentar varios cuestionarios y en diferentes momentos a lo largo del estudio.

La reducción de las escalas del cuestionario WOMAC ya ha sido abordada. Algunos autores han trabajado con las dimensiones de dolor y capacidad funcional<sup>9,10</sup>, mientras que otros han estudiado la dimensión de capacidad funcional<sup>11,12</sup>. Entre estas últimas, se elaboró una escala reducida de siete ítems<sup>12</sup>.

El objetivo del presente estudio fue evaluar la validez, la fiabilidad y la sensibilidad al cambio de la versión española de la dimensión reducida de siete ítems de capacidad funcional del WOMAC desarrollada originalmente por Whitehouse et al<sup>12</sup>.

## Métodos

Estudio prospectivo con seguimiento de 1 año realizado en 15 hospitales, tres de Andalucía, tres de las Islas Canarias y nueve del País Vasco, cuyos comités de ética e investigación clínica aprobaron el estudio.

Los pacientes elegibles para el estudio fueron aquellos incluidos en lista de espera para recibir una prótesis total de cadera o de rodilla por artrosis primaria, entre marzo de 2005 y diciembre de 2006.

Se excluyeron los pacientes con enfermedad psiquiátrica, ya que ésta podía dificultar la cumplimentación de los cuestionarios incluidos.

Los cuestionarios se enviaron 1 mes antes de la intervención quirúrgica y a los 3, 6 y 12 meses después de realizarla. El tamaño muestral del análisis secundario se basó en los pacientes que respondieron a los cuatro envíos.

## Cuestionarios

Los pacientes cumplimentaron el cuestionario WOMAC, que es un instrumento específico de CVRS<sup>5</sup>. Es una escala multidimensional compuesta por 24 ítems, en tres dimensiones: dolor (5 ítems), rigidez (2 ítems) y capacidad funcional (17 ítems). La versión original del cuestionario ha sido traducida<sup>13</sup> y validada en español<sup>14</sup>.

Utilizamos la versión tipo Likert, con cinco opciones de respuesta. La pregunta sobre capacidad funcional se refiere al grado de dificultad para realizar 17 actividades, y las opciones de respuesta son en todas ellas iguales: ninguna, poca, bastante, mucha y muchísima dificultad, con una puntuación que va de 0 a 4, respectivamente. Luego se suman los puntos obtenidos y se estandarizan de 0 a 100 (de mejor a peor capacidad). Los ítems incluidos en la escala reducida son subir escaleras, levantarse después de estar sentado, andar en un terreno llano, entrar o salir del coche, ponerse los calcetines, levantarse de la cama y estar sentado.

Asimismo, se remitió el cuestionario genérico de CVRS Euro-QoL 5D (EQ-5D)<sup>15</sup>. Se utilizó la descripción del estado de salud que tiene cinco dimensiones (movilidad, cuidado personal, actividades cotidianas, dolor/malestar y ansiedad/depresión), con tres niveles de gravedad (sin problemas, algunos/moderados y muchos problemas). Este cuestionario ha sido traducido y validado en español<sup>16</sup>.

## Análisis estadístico

Los datos descriptivos se presentan en forma de frecuencias con porcentajes y medias junto a las desviaciones estándar (DE). Para la comparación de variables cualitativas usamos la prueba de ji al cuadrado, y la *t* de Student para variables cuantitativas.

La validez se refiere a que un instrumento mide realmente lo que se supone que tiene que medir. En este caso hemos estudiado la validez de criterio y de constructo. La validez de criterio es la valoración del nuevo instrumento frente el verdadero valor. Este tipo de validez se ha estudiado básicamente en casos como el presente, el estudio de versiones reducidas empleando el original como estándar<sup>17</sup>. Se analizó mediante el coeficiente de correlación rho de Spearman entre ambas escalas en los cuatro tiempos del estudio.

Para evaluar la validez de constructo hemos utilizado varios análisis. Al igual que en un estudio de validación de la misma escala<sup>18</sup>, usamos el coeficiente de concordancia denominado correlación de Lin<sup>19</sup> entre ambas escalas.

Hemos utilizado el gráfico de Bland y Altman<sup>20</sup>, que consiste en representar las diferencias entre las dos mediciones, escalas reducida y completa, frente a la media de ambas, junto a los límites de acuerdo del 95%. Es útil para ver diferencias sistemáticas y si las diferencias entre las puntuaciones dependen del nivel de éstas.

Además, utilizamos el análisis de grupos conocidos. Esto significa que ciertos grupos de pacientes deberían obtener puntuaciones diferentes basadas en alguna característica. En nuestro caso usamos la pregunta realizada en la medición basal del EQ-5D sobre dificultades para realizar actividades cotidianas. La hipótesis es que aquellos pacientes con mayor dificultad obtendrían peores puntuaciones que los que tienen menos problemas para realizar las actividades diarias. Se utilizó el análisis de la varianza con la prueba de Scheffé para comparaciones múltiples.

Se analizó la unidimensionalidad de los siete ítems de la versión reducida. Para ello hemos utilizado el modelo Rasch ajustado al modelo teórico denominado *rating scale model*<sup>21</sup>. Los estadísticos basados en el ajuste de la ji al cuadrado, conocidos como INFIT y OUTFIT, se emplearon para comprobar la validez de los ítems. Se consideró que un ítem era adecuado cuando sus valores se encontraban entre 0,7 y 1,3<sup>22</sup>. Este modelo aporta datos sobre la dificultad de los diferentes ítems. Los más difíciles se indican por los valores negativos de sus logit. El logit medio de un grupo de ítems que se ajustan al modelo Rasch debería ser 0 y su DE debería ser 1<sup>10</sup>. Se presentan los índices de separación de personas e ítems con sus correspondientes grados de fiabilidad. Los índices han de ser mayores de 2 y la fiabilidad de 0,80.

Para valorar las posibles diferencias en el grado de dificultad para realizar las actividades de los ítems se ha analizado el funcionamiento diferencial del ítem (DIF), según la articulación afectada. Para ello se analizaron de forma conjunta los pacientes

**Tabla 1**  
Datos descriptivos de la muestra

	Prótesis total de rodilla (n = 263)	Prótesis total de cadera (n = 196)	p
Sexo, mujeres (%)	174 (66,2)	84 (42,8)	< 0,001
Edad, media (DE)	71,7 (6,5)	65,6 (12,2)	< 0,001
IMC, media (DE)	30,3 (5,1)	28,0 (4,0)	< 0,001
<i>Escala capacidad funcional WOMAC<sup>a</sup> completa, media (DE)</i>			
Basal	58,8 (18,2)	63,9 (16,9)	0,003
3 meses	31,3 (17,9)	30,3 (16,7)	0,50
6 meses	28,9 (18,8)	25,5 (16,5)	0,04
1 año	28,3 (21,2)	22,5 (17,5)	0,002
<i>Escala capacidad funcional WOMAC<sup>a</sup> reducida, media (DE)</i>			
Basal	57,9 (18,2)	64,5 (16,9)	< 0,001
3 meses	29,9 (17,4)	29,5 (16,5)	0,8
6 meses	28,1 (19,2)	25,2 (16,1)	0,08
1 año	27,0 (20,9)	21,9 (21,0)	0,006

IMC: índice de masa corporal; DE: desviación estándar; WOMAC: Western Ontario McMaster Universities Osteoarthritis Index.

<sup>a</sup> Valores estandarizados de 0 (mejor) a 100 (peor).

intervenidos de prótesis total de cadera y de prótesis total de rodilla. Las diferencias por encima de 0,64 se consideraron clínicamente importantes<sup>23</sup>.

#### Fiabilidad

Medimos la consistencia interna de la escala por medio del coeficiente  $\alpha$  de Cronbach<sup>24</sup>. Valores por encima de 0,7 se consideran aceptables, entre 0,7 y 0,9 son buenos, y por encima de 0,9 son excelentes<sup>25</sup>.

#### Sensibilidad al cambio

La sensibilidad al cambio se determinó por la respuesta media estandarizada (RME), que se calcula dividiendo el cambio medio entre dos mediciones por su DE. Por lo tanto, si el cambio tiene una gran variabilidad respecto a su media, se obtendrá un valor bajo de RME<sup>26</sup>. En nuestro caso se calculó la RME respecto a la medición basal. Los valores de 0,20, 0,50 y 0,80 indican escasa, moderada y alta sensibilidad al cambio<sup>27</sup>.

Los análisis estadísticos se realizaron con los programas SPSS versión 17.0 y WINSTEP versión 3.71

**Tabla 2**  
Validez de criterio y de constructo de la versión completa (17 ítems) y reducida (7 ítems) de la dimensión de capacidad funcional del WOMAC

	EuroQol-5D, actividades diarias			p
	Sin problemas <sup>a</sup> , media (DE)	Algunos problemas <sup>b</sup> , media (DE)	Muchos problemas <sup>c</sup> , media (DE)	
<b>Prótesis total de rodilla</b>				
<i>Capacidad funcional WOMAC completa<sup>d</sup></i>				
Basal	45,1 (17,2) <sup>b,c</sup>	60,9 (16,6) <sup>a</sup>	68,3 (18,2) <sup>a</sup>	< 0,001
1 año	15,9 (13,4) <sup>b,c</sup>	41,2 (18,6) <sup>a,c</sup>	67,5 (17,6) <sup>a,b</sup>	< 0,001
<i>Capacidad funcional WOMAC reducida</i>				
Basal	44,8 (17,2) <sup>b,c</sup>	59,7 (17,0) <sup>a</sup>	67,5 (17,1) <sup>a</sup>	< 0,001
1 año	15,2 (13,4) <sup>b,c</sup>	39,6 (19,2) <sup>a,c</sup>	65,0 (12,5) <sup>a,b</sup>	< 0,001
<b>Prótesis total de cadera</b>				
<i>Capacidad funcional WOMAC completa<sup>d</sup></i>				
Basal	44,0 (13,4) <sup>b,c</sup>	63,7 (15,6) <sup>a,c</sup>	75,7 (15,8) <sup>a,b</sup>	< 0,001
1 año	14,0 (10,6) <sup>b,c</sup>	32,8 (16,3) <sup>a,c</sup>	55,0 (31,5) <sup>a,b</sup>	< 0,001
<i>Capacidad funcional WOMAC reducida</i>				
Basal	46,7 (13,3) <sup>b,c</sup>	64,3 (16,1) <sup>a,c</sup>	75,4 (15,7) <sup>a,b</sup>	< 0,001
1 año	14,1 (11,2) <sup>b,c</sup>	31,2 (16,4) <sup>a,c</sup>	52,1 (32,3) <sup>a,b</sup>	< 0,001

DE: desviación estándar; WOMAC: Western Ontario McMaster Universities Osteoarthritis Index.

<sup>a,b,c</sup> Las letras voladas indican diferencias entre grupos. Así, si aparece en el grupo<sup>(a)</sup> las letras<sup>(b,c)</sup> quieren decir que hay diferencias estadísticamente significativas entre los tres grupos.

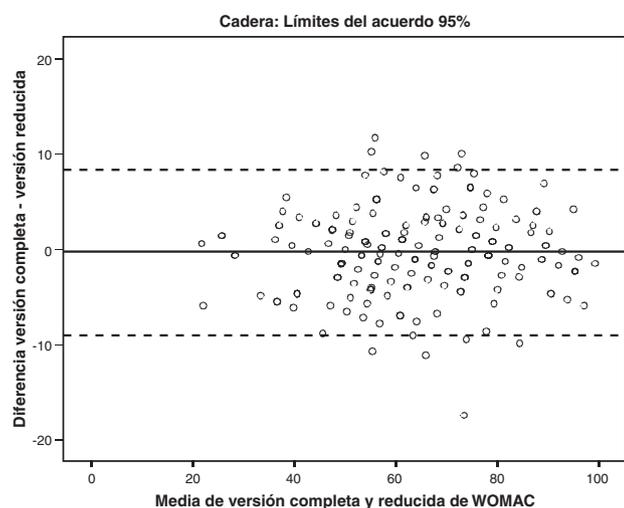
<sup>d</sup> Valores estandarizados de 0 (mejor) a 100 (peor).

## Resultados

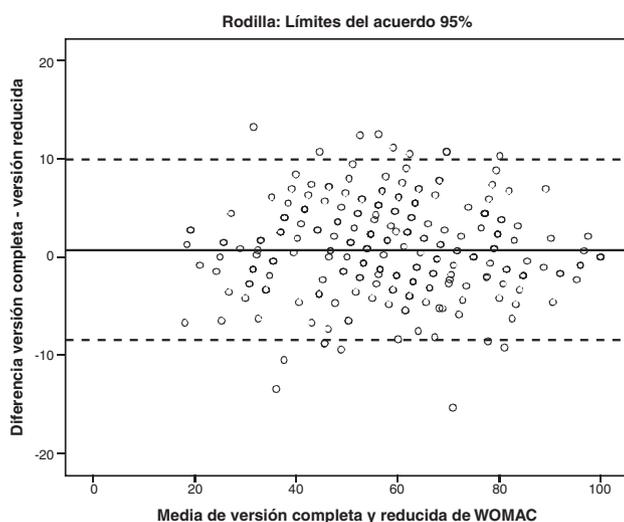
La muestra del estudio principal incluyó mediciones basales del cuestionario WOMAC en 1660 pacientes; a los 3 meses obtuvimos respuestas de 899, a los 6 meses de 818 y al año de 883 pacientes. La muestra del presente trabajo consistió en 459 pacientes que cumplieron los cuestionarios en las cuatro mediciones. La distribución de pacientes incluidos por comunidad autónoma fue de 65 en Canarias (14,16%), 66 en Andalucía (14,38%) y 328 en el País Vasco (71,46%). Hubo 263 prótesis totales de rodilla (57,3%) y 196 de cadera (42,7%). Como muestra la **tabla 1**, entre los pacientes con prótesis total de rodilla había más mujeres ( $p < 0,001$ ), eran de edad avanzada ( $p < 0,001$ ) y tenían un mayor índice de masa corporal ( $p < 0,001$ ). En ambas articulaciones, tanto las puntuaciones medias como las DE de la escala completa de capacidad funcional y de la reducida fueron similares en todos los tiempos medidos ( $p > 0,05$  en todos los casos). Las diferencias entre las medias de las puntuaciones fueron de alrededor de 1 punto.

En ambas articulaciones, y en las cuatro mediciones, los valores de los coeficientes rho de Spearman y de concordancia de Lin entre la versión completa y la reducida estuvieron entre 0,96 y 0,98. La **tabla 2** muestra las comparaciones entre grupos conocidos en dos momentos: basal y al año tras cirugía. Los pacientes que no tenían problemas en sus actividades cotidianas tuvieron la mejor puntuación, y los que tenían muchos problemas obtuvieron las peores ( $p < 0,001$ ); aquellos que presentaron una situación intermedia, tenían puntuaciones intermedias. Este gradiente en las puntuaciones se observó en ambas articulaciones y en ambas mediciones. Del mismo modo, los valores fueron similares para la versión completa y la reducida. En general, las diferencias observadas no fueron mayores de 1,5 puntos entre ambas escalas. Las diferencias encontradas en el grupo de pacientes intervenidos de prótesis total de cadera fueron estadísticamente significativas entre los tres grupos de dificultad ( $p < 0,001$ ) y en los dos momentos. En el caso de la prótesis total de rodilla, las diferencias fueron muy similares a las de la prótesis total de cadera.

Los gráficos de Bland y Altman muestran que tanto en la prótesis total de cadera (**fig. 1**) como en la de rodilla (**fig. 2**) la variabilidad es aleatoria y uniforme en todo el intervalo de valores. En el caso de la prótesis total de cadera, la versión reducida puntúa más alto, con una diferencia media de -0,25 (DE: 4,4), al revés que en el caso de la prótesis total de rodilla, en la cual la diferencia es de 0,82 (DE: 4,7). Los límites de acuerdo muestran que el 95% de las diferencias están



**Figura 1.** Gráfico de Bland y Altman para las diferencias en las puntuaciones de la versión completa (17 ítems) y reducida (7 ítems) de la dimensión de capacidad funcional del WOMAC (prótesis total de cadera).



**Figura 2.** Gráfico de Bland y Altman para las diferencias en las puntuaciones de la versión completa (17 ítems) y reducida (7 ítems) de la dimensión de capacidad funcional del WOMAC (prótesis total de rodilla).

entre -8,98 y 8,48 puntos en el caso de la cadera y entre -8,38 y 10,02 en la rodilla.

Los datos de la **tabla 3**, pertenecientes al análisis Rasch, nos muestran que los valores INFIT y OUTFIT apoyan la unidimensionalidad, salvo el ítem de ponerse los calcetines. Asimismo, los supuestos sobre la media y la desviación estándar se cumplen. El índice de separación de personas fue de 2,49, y el de ítems de 15,17,

**Tabla 3**  
Análisis Rasch de la dimensión reducida (7 ítems) de capacidad funcional del WOMAC

Ítem de dificultad	Descripción del ítem	Logit (EE)	INFIT	OUTFIT	Rango basado en el logit
7	Estar sentado	2,08 (0,07)	1,21	1,21	1
3	Andar en llano	0,74 (0,07)	0,87	0,91	2
6	Levantarse de la cama	0,44 (0,07)	0,87	0,86	3
2	Levantarse de una silla	-0,68 (0,07)	0,80	0,80	4
5	Ponerse calcetines	-0,74 (0,07)	1,38	1,31	5
1	Subir escaleras	-0,79 (0,07)	0,95	1,02	6
4	Entrar/salir coche	-1,05 (0,07)	0,86	0,83	7
	Media (DE)	0,0 (1,06)			

WOMAC: Western Ontario McMaster Universities Osteoarthritis Index; EE: error estándar; DE: desviación estándar.

**Tabla 4**

Fiabilidad y sensibilidad al cambio de la versión completa (17 ítems) y reducida (7 ítems) de la dimensión de capacidad funcional del WOMAC

	Basal	3 meses	6 meses	1 año
<b>Prótesis total de rodilla</b>				
<i>Escala capacidad funcional WOMAC</i>				
Completa RME		1,28	1,38	1,27
Reducida RME		1,29	1,29	1,27
Completa $\alpha$ de Cronbach	0,94	0,95	0,96	0,97
Reducida $\alpha$ de Cronbach	0,86	0,89	0,92	0,94
<b>Prótesis total de cadera</b>				
<i>Escala capacidad funcional WOMAC</i>				
Completa RME		1,66	1,85	1,98
Reducida RME		1,72	1,85	2,03
Completa $\alpha$ de Cronbach	0,94	0,95	0,96	0,96
Reducida $\alpha$ de Cronbach	0,85	0,88	0,89	0,90

WOMAC: Western Ontario McMaster Universities Osteoarthritis Index; RME: respuesta media estandarizada.

correspondiendo a valores de fiabilidad superiores a 0,80 en ambos casos.

Según el análisis del funcionamiento diferencial, el ítem de dificultad para ponerse los calcetines resultó ser más difícil en el grupo de pacientes intervenidos de prótesis total de cadera, con una diferencia de 1,41 logit. La **tabla 3** presenta las diferentes actividades ordenadas según su dificultad. En el conjunto de los pacientes, la actividad más sencilla de realizar es estar sentado (logit 2,08), mientras que la de mayor dificultad es entrar o salir de un coche (logit -1,05).

Finalmente, la **tabla 4** muestra los datos de fiabilidad y sensibilidad al cambio de ambas versiones. En los pacientes con prótesis total de rodilla, los valores del coeficiente  $\alpha$  de Cronbach fueron altos para la versión completa, de 0,94 a 0,97 según el momento en que se midió. En la versión reducida estos valores fueron inferiores, pero todos ellos estaban en el intervalo de 0,86 a 0,94. En el caso de la prótesis total de cadera, los valores de la versión completa van de 0,94 a 0,96, mientras que los de la versión reducida oscilan entre 0,85 y 0,90. Estos valores están por encima del recomendado de 0,70.

En cuanto a la sensibilidad al cambio, en los pacientes con prótesis total de rodilla la versión completa obtuvo valores ligeramente diferentes según el momento de la medición. Tanto en la versión completa como en la reducida, el tamaño del efecto, superior a 1,26, puede considerarse alto en todos los casos. En los pacientes intervenidos de prótesis total de cadera, estos valores son aún mayores y similares entre ambas escalas, y demuestran una elevada sensibilidad al cambio, que aumentaba con el tiempo.

## Discusión

Los datos de este estudio prospectivo realizado en pacientes con prótesis total de rodilla o de cadera, con un seguimiento de 1 año, aportan pruebas que apoyan el adecuado funcionamiento de la versión reducida de la escala de capacidad funcional del cuestionario

WOMAC. En este estudio hemos utilizado la versión reducida de siete ítems, pues ya había demostrado en su versión original que era una escala válida, fiable y sensible al cambio en los pacientes ambulatorios con artrosis de rodilla<sup>18</sup> y en los valorados antes de una artroplastia<sup>28</sup>.

El cuestionario WOMAC es ampliamente utilizado tanto en estudios de investigación en cirugía ortopédica o procesos reumatológicos como en la práctica clínica habitual<sup>11,6</sup>. Uno de los criterios de valoración en el uso de instrumentos de CVRS es la carga que puede suponer su cumplimentación<sup>29</sup>. En los estudios epidemiológicos, la carga para los pacientes podría medirse en términos del número y la extensión de los cuestionarios a cumplimentar. El objetivo en la práctica clínica sería recoger información para evaluar la respuesta al tratamiento en términos útiles para el paciente y para el médico, con el menor esfuerzo. En ambos casos, si se consigue la misma información con menor carga para el paciente mediante el uso de versiones reducidas, como esta escala de capacidad funcional del WOMAC, el instrumento resultará útil<sup>30</sup>.

Los autores de la versión reducida de siete ítems<sup>12</sup> utilizaron dos enfoques para llevar a cabo el proceso. En primer lugar, bajo un punto de vista clínico con la participación de diferentes expertos, y posteriormente mediante análisis psicométrico. Nuestros resultados son similares a los obtenidos en el estudio original, tanto en lo relativo a las puntuaciones basales como a 3 y 12 meses en ambas versiones de la escala. Esto significa, además, que las puntuaciones de nuestras escalas, la completa y la reducida, fueron similares en todos los momentos de medición, al igual que ocurrió en la versión original.

En comparación con la versión completa, los coeficientes de correlación de Spearman en las cuatro mediciones de nuestro estudio son buenos, con valores entre 0,96 y 0,98 para ambas articulaciones. Estos valores son similares a los obtenidos en la validación original de la escala<sup>12</sup>.

Respecto a los datos sobre la validez de constructo, podemos decir que los coeficientes de Lin muestran valores muy buenos, tanto en los pacientes con prótesis total de rodilla como de cadera. Estas cifras se mantienen de forma constante a lo largo de todos los tiempos medidos. Los valores obtenidos son similares a los hallados en el estudio de validación llevado a cabo en pacientes ambulatorios<sup>18</sup>. Observando los gráficos de Bland y Altman podemos concluir que no hay diferencias sistemáticas, así como que las diferencias entre ambas escalas no dependen de la gravedad en ambos tipos de pacientes.

La escala reducida diferencia adecuadamente entre los grupos de gravedad medidos por una pregunta del EQ-5D, tanto en prótesis total de rodilla como de cadera. Se ve cómo se establece un gradiente de puntuación de menor a mayor dificultad para realizar las actividades diarias. Además, salvo en la medición basal llevada a cabo en pacientes con prótesis total de rodilla, las diferencias encontradas se establecen entre los tres grupos estudiados.

Una característica básica para evaluar la validez de constructo es la unidimensionalidad de la escala. El análisis Rasch es un método para poder comprobarla. En la literatura existen diferentes trabajos que estudian este aspecto, pero sus resultados no coinciden, pues algunos concluyen que la dimensión de capacidad funcional no se ajusta al modelo Rasch<sup>9,10,31</sup> y otros más recientes afirman la unidimensionalidad de la escala completa<sup>32</sup>, atribuyendo los anteriores resultados negativos a un sesgo de función diferencial (DIF) de algunos ítems. En nuestro caso, el análisis apoya la unidimensionalidad de la escala reducida, y además muestra que el ítem sobre la dificultad para ponerse los calcetines funciona diferente según sea la articulación estudiada (en los pacientes con prótesis total de cadera este ítem es el que presenta mayores dificultades).

Respecto a la fiabilidad, como era de esperar por el número de ítems incluidos, los valores de la versión reducida son un poco inferiores a los de la versión completa, pero siempre superiores a 0,85, por lo que pueden considerarse como buenos. Nuestros valores del coeficiente  $\alpha$  de Cronbach son muy similares a los obtenidos en los estudios de validación de la versión original reducida<sup>12,18</sup>.

Las cifras de RME obtenidas indican que ambas versiones tienen una similar sensibilidad al cambio. En ambas, los valores son más altos en los pacientes con prótesis total de cadera. Las cifras presentadas en el estudio original<sup>12</sup> son similares en la medición de los 3 meses: alrededor de 1,3 en los pacientes con prótesis total de rodilla. Por otro lado, aunque el estudio realizado en pacientes ambulatorios<sup>18</sup> aporta datos sobre la RME, no pueden compararse, ya que tienen que presentar respuestas menores por el tratamiento aplicado.

Finalmente, otros autores desarrollaron una escala reducida, de ocho ítems, a partir de las opiniones de pacientes y de reumatólogos<sup>11</sup>. En este estudio, así como en otro que llevó a cabo una validación de dicha escala<sup>33</sup>, los datos sobre la validez de constructo medida por la correlación entre ambas escalas ( $r=0,95$ ) fueron similares a los de nuestro estudio. Asimismo, los valores de consistencia interna también fueron similares, con valores  $\alpha$  de Cronbach de 0,84 y 0,93 para las escalas reducida y completa, respectivamente. La sensibilidad al cambio, sin embargo, no es comparable porque los dos estudios se llevaron a cabo con pacientes ambulatorios.

Debería tenerse en cuenta que la validación de la presente versión de siete ítems sólo sería aplicable a aquellos pacientes con artroplastia de cadera o rodilla, ya que los valores de validez, fiabilidad o sensibilidad al cambio pueden cambiar en otras circunstancias clínicas, como pacientes en rehabilitación o bajo tratamiento médico.

En conclusión, con los resultados de este estudio hemos añadido elementos de prueba que apoyan el uso de la versión corta de la escala de capacidad funcional del WOMAC en los pacientes que van a ser intervenidos de prótesis total de cadera o de rodilla.

### ¿Qué se sabe sobre el tema?

En los pacientes con artrosis de miembros inferiores se han utilizado diferentes instrumentos específicos para medir el dolor, la capacidad funcional y la calidad de vida relacionada con la salud. Uno de los más utilizados es el Western Ontario McMaster Universities Osteoarthritis Index (WOMAC), que está considerado como uno de los mejores cuestionarios desde el punto de vista de sus propiedades psicométricas. Este cuestionario tiene tres dimensiones que miden dolor, rigidez y capacidad funcional. La dimensión de capacidad funcional es la más larga de las tres (en su versión original, 17 preguntas). Este hecho puede causar una baja tasa de respuesta, sobre todo en estudios en los cuales haya una importante carga para el paciente en términos del número de preguntas a responder y si las mediciones se llevan a cabo en varios momentos.

### ¿Qué añade el estudio realizado a la literatura?

El presente estudio aporta datos sobre la validez, la fiabilidad y la sensibilidad al cambio de una versión reducida de siete ítems sobre la dimensión de capacidad funcional del cuestionario WOMAC. Es posible que su uso incremente las tasas de respuesta en estudios longitudinales, sobre todo si implican varias mediciones en el seguimiento de estos pacientes. Asimismo, puede favorecer su uso en la práctica clínica habitual al disminuir el tiempo de cumplimentación.

## Contribuciones de autoría

A. Escobar, J.M. Quintana, L. García Pérez y C. Herrera-Espiñeira concibieron el estudio original, contribuyeron en la organización y recogida de datos del estudio y en la redacción del artículo. K. Vrotsou y A. Bilbao realizaron el análisis y la interpretación de los datos, así como la revisión crítica del artículo.

## Financiación

El estudio original fue financiado como proyecto multicéntrico por el Fondo de Investigaciones Sanitarias (PI 04/0938, PI 04/0542, y PI 04/2577) y por el Departamento de Sanidad del Gobierno Vasco (200411012).

## Conflicto de intereses

Ninguno.

## Agradecimientos

Al personal de los diferentes servicios implicados, unidades de investigación y de calidad, así como a las unidades de documentación clínica de los hospitales participantes y finalmente a los pacientes por su colaboración.

## Bibliografía

- Ethgen O, Bruyere O, Richy F, et al. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. *J Bone Joint Surg Am.* 2004;86-A:963-74.
- Escobar A, Quintana JM, Bilbao A, et al. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. *Osteoarthritis Cartilage.* 2007;15:273-80.
- Quintana JM, Escobar A, Arostegui I, et al. Health-related quality of life and appropriateness of knee or hip joint replacement. *Arch Intern Med.* 2006;166:220-6.
- Veenhof C, Bijlsma JW, van den Ende CH, et al. Psychometric evaluation of osteoarthritis questionnaires: a systematic review of the literature. *Arthritis Rheum.* 2006;55:480-92.
- Bellamy N, Buchanan WW, Goldsmith CH, et al. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol.* 1988;15:1833-40.
- McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. *Arthritis Rheum.* 2001;45:453-61.
- Bellamy N, Kirwan J, Boers M, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol.* 1997;24:799-802.
- Hawker G, Melfi C, Paul J, et al. Comparison of a generic (SF-36) and a disease specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery. *J Rheumatol.* 1995;22:1193-6.
- Rothenfluh DA, Reedwisch D, Muller U, et al. Construct validity of a 12-item WOMAC for assessment of femoro-acetabular impingement and osteoarthritis of the hip. *Osteoarthritis Cartilage.* 2008;16:1032-8.
- Davis AM, Badley EM, Beaton DE, et al. Rasch analysis of the Western Ontario McMaster (WOMAC) Osteoarthritis Index: results from community and arthroplasty samples. *J Clin Epidemiol.* 2003;56:1076-83.
- Tubach F, Baron G, Falissard B, et al. Using patients' and rheumatologists' opinions to specify a short form of the WOMAC function subscale. *Ann Rheum Dis.* 2005;64:75-9.
- Whitehouse SL, Lingard EA, Katz JN, et al. Development and testing of a reduced WOMAC function scale. *J Bone Joint Surg Br.* 2003;85:706-11.
- Batlle-Gualda E, Esteve-Vives J, Piera MC, et al. Traducción y adaptación al español del cuestionario WOMAC específico para artrosis de rodilla y cadera. *Rev Esp Reumatol.* 1999;26:38-45.
- Escobar A, Quintana JM, Bilbao A, et al. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster Universities Osteoarthritis Index. *Clin Rheumatol.* 2002;21:466-71.
- Rabin R, de CF. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med.* 2001;33:337-43.
- Badia X, Roset M, Montserrat S, et al. La versión española del EuroQol: descripción y aplicaciones. *Med Clin (Barc).* 1999;112 Supl 1:79-85.
- Valderas JM, Ferrer M, Alonso J. Instrumentos de medida de calidad de vida relacionada con la salud y de otros resultados percibidos por los pacientes. *Med Clin (Barc).* 2005;125 Supl 1:56-60.
- Yang KG, Raijmakers NJ, Verbout AJ, et al. Validation of the short-form WOMAC function scale for the evaluation of osteoarthritis of the knee. *J Bone Joint Surg Br.* 2007;89:50-6.
- Lin LL. A concordance correlation coefficient to evaluate reproducibility. *Bio-metrics.* 1989;45:255-68.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307-10.
- Baghaei P. A comparison of three polychotomus Rasch models for super-item analysis. *Psychological Test and Assessment Modeling.* 2010;52:313-22.
- Bond TG, Fox CM. Applying the Rasch model. Fundamental measurement in the human sciences. 2nd ed. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc; 2007.
- Scott NW, Fayers PM, Aaronson NK, et al. A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *J Clin Epidemiol.* 2009;62:288-95.
- Cronbach LJ. Coefficient alpha and the internal structure of test. *Psychometrika.* 1951;16:297-334.
- Nunnally JC, Bernstein IH. Psychometric theory. 3rd. ed. New York: McGraw-Hill; 1994.
- García de Yébenes Prous MJ, Rodríguez Salvanés F, Carmona Ortells L. Sensibilidad al cambio de las medidas de desenlace. *Reumatol Clin.* 2008;4:240-7.
- Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol.* 2000;53:459-68.
- Whitehouse SL, Crawford RW, Learmonth ID. Validation for the reduced Western Ontario and McMaster Universities Osteoarthritis Index function scale. *J Orthop Surg (Hong Kong).* 2008;16:50-3.
- Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res.* 2002;11:193-205.
- Singh J, Sloan JA, Johanson NA. Challenges with health-related quality of life assessment in arthroplasty patients: problems and solutions. *J Am Acad Orthop Surg.* 2010;18:72-82.
- Wolfe F, Kong SX. Rasch analysis of the Western Ontario MacMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Ann Rheum Dis.* 1999;58:563-8.
- Kersten P, White PJ, Tennant A. The visual analogue WOMAC 3.0 scale - internal validity and responsiveness of the VAS version. *BMC Musculoskelet Disord.* 2010;11:80.
- Baron G, Tubach F, Ravaut P, et al. Validation of a short form of the Western Ontario and McMaster Universities Osteoarthritis Index function subscale in hip and knee osteoarthritis. *Arthritis Rheum.* 2007;57:633-8.